# TITLE OF THE INVENTION

# INFORMATION PROCESSING APPARATUS, INFORMATION PROCESSING METHOD AND PROGRAM, AND STORAGE MEDIUM

## FIELD OF THE INVENTION

The present invention relates to a process for inputting English pronunciation symbols.

## BACKGROUND OF THE INVENTION

Upon developing a speech synthesis English dictionary or creating English phonetic text, an English pronunciation symbol string must be input. However, English pronunciation symbols cannot be intuitively input unlike Japanese reading.

As conventional methods of inputting English pronunciation symbols (about 40 symbols), a method of registering pronunciation symbols as external characters and selecting them from an external character symbol table, a method of setting each of pronunciation symbols in correspondence with one or two alphabets and inputting symbols like normal text, and the like are known (for example, see Japanese Patent Laid-Open No. 7-78133).

However, with the method of registering pronunciation symbols as external characters, the user must display the external character symbol table and select a symbol from it every time he or she inputs one

pronunciation symbol, resulting in an inefficient input process. Also, since external characters are used, compatibility with other systems is poor.

Furthermore, with the method of setting each pronunciation symbol in correspondence with one or two alphabets, it is difficult for the user to intuitively recognize the correspondence between an alphabet string and pronunciation symbol and to accurately input symbols.

## SUMMARY OF THE INVENTION

The present invention has been made in consideration of the above problems, and has as its object to provide a processing technique that allows the user to efficiently and accurately input pronunciation symbols.

In order to achieve the above object, an information processing apparatus according to the present invention comprises the following arrangement. That is, an information processing apparatus for inputting a pronunciation symbol corresponding to an English notation, comprising: pronunciation symbol information holding means for holding pronunciation symbol information indicating a relationship between a predetermined alphabet and a pronunciation symbol that starts from the predetermined alphabet; pronunciation symbol statistical information holding means for

holding statistical information associated with a probability of occurrence of each pronunciation symbol immediately after a predetermined pronunciation symbol; display means for extracting pronunciation symbols corresponding to an input alphabet from the pronunciation symbol information, and displaying the extracted pronunciation symbols while sorting them on the basis of the statistical information; and determination means for determining a pronunciation symbol corresponding to the English notation from the displayed pronunciation symbols.

The information processing apparatus of the present invention allows the user to efficiently and accurately input pronunciation symbols.

Other features and advantages of the present invention will be apparent from the following description taken in conjunction with the accompanying drawings, in which like reference characters designate the same or similar parts throughout the figures thereof.

## BRIEF DESCRIPTION OF THE DRAWINGS

The accompanying drawings, which are incorporated in and constitute a part of the specification, illustrate embodiments of the invention and, together with the description, serve to explain the principles of the invention.

Fig. 1 is a block diagram showing the arrangement of an information processing apparatus according to an embodiment of the present invention;

Fig. 2 is a flow chart showing the processing sequence of the information processing apparatus according to the embodiment of the present invention;

Fig. 3 shows a pronunciation symbol table 105 of the information processing apparatus according to the embodiment of the present invention;

Fig. 4 shows an associative pronunciation symbol table 106 of the information processing apparatus according to the embodiment of the present invention;

Fig. 5 shows pronunciation symbol statistical information 107 of the information processing apparatus according to the embodiment of the present invention;

Fig. 6 shows pronunciation symbol image data 108 of the information processing apparatus according to the embodiment of the present invention;

Fig. 7 shows pronunciation symbol auxiliary data 109 of the information processing apparatus according to the embodiment of the present invention;

Fig. 8 shows an edit result database 118 of the information processing apparatus according to the embodiment of the present invention; and

Fig. 9 shows an edit process of pronunciation symbols by the information processing apparatus according to the embodiment of the present invention.

## DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

Preferred embodiments of the present invention will now be described in detail in accordance with the accompanying drawings.

Fig. 1 is a block diagram showing the arrangement of an information processing apparatus according to an embodiment of the present invention.

Reference numeral 101 denotes a notation processing unit that executes a process associated with English notations to which pronunciation symbols are to be given.

Reference numeral 102 denotes a pronunciation symbol candidate processing unit that executes a process associated with pronunciation symbol candidates. Reference numeral 103 denotes a pronunciation symbol candidate holding unit that holds pronunciation symbol candidates. Reference numeral 104 denotes a pronunciation symbol candidate presentation unit that presents pronunciation symbol candidates. Reference numeral 105 denotes a pronunciation symbol table that stores alphabets and pronunciation symbols each of which has a corresponding alphabet as its first character. Fig. 3 shows an example of the pronunciation symbol table.

Reference numeral 106 denotes an associative pronunciation symbol table that stores alphabets, and

pronunciation symbols each of which is associable as the pronunciation of a given alphabet when that alphabet forms a part of an arbitrary English notion. Fig. 4 shows an example of the associative pronunciation symbol table. For example, pronunciation symbols of an English notation "able" is "EY1 B AH0 L," and "EY" is associable as the pronunciation of alphabet "a."

Reference numeral 107 denotes pronunciation symbol statistical information used to determine a presentation order of pronunciation symbol candidates. Fig. 5 shows an example of the pronunciation symbol statistical information. In this case, a statistical value is generated by multiplying by -1 the logarithm of the probability of occurrence of a pronunciation symbol of interest immediately after a forward pronunciation symbol, and normalizing the product to an integer by multiplying that product by an appropriate value. Symbol Φ indicates a case wherein no forward pronunciation symbol is present ( i.e., a case wherein the pronunciation symbol of interest is located at the head of an English notation). The probability of occurrence of a pronunciation symbol of interest immediately after a forward pronunciation symbol can be generated based on a dictionary and the like.

Reference numeral 108 denotes pronunciation symbol image data as pairs of pronunciation symbols

expressed by alphabets and image symbols (symbols generally used in a dictionary and the like) corresponding to these pronunciation symbols. Fig. 6 shows an example of the pronunciation symbol image data. Reference numeral 109 denotes pronunciation symbol auxiliary data as pairs of pronunciation symbols expressed by alphabets, and auxiliary data of these pronunciation symbols. Fig. 7 shows an example of the pronunciation symbol auxiliary data. "odd: AA D" indicates that the pronunciation symbol "AA" is a pronunciation of "AA" of "odd."

Reference numeral 110 denotes a key input processing unit that processes key operations input by the user upon editing pronunciation symbols. Reference numeral 111 denotes an input alphabet holding unit that holds alphabets input by the user.

Reference numeral 112 denotes an input mode change unit that changes an input mode between two input modes (i.e., a direct input mode and associative input mode). In the direct input mode, the user directly inputs and edits the first alphabet of a pronunciation symbol. In the associative input mode, the user inputs and edits some alphabets of an English notation to which pronunciation symbols are to be given. Reference numeral 113 denotes an input mode holding unit that holds the current input mode.

Reference numeral 114 denotes a pronunciation

symbol determination unit that processes a pronunciation symbol determination operation. Reference numeral 115 denotes a pronunciation symbol speech generation unit for generating speech of pronunciation symbols. Reference numeral 116 denotes a phonemic symbol dictionary as acoustic data used to generate speech of pronunciation symbols. Reference numeral 117 denotes an edit result save unit that saves the edit results of pronunciation symbols. Reference numeral 118 denotes an edit result database that holds the edit results of pronunciation symbols. Fig. 8 shows an example of the edit result database. In this case, the database holds pairs of English notations and pronunciation symbols.

Fig. 2 is a flow chart showing the processing sequence in the information processing apparatus according to the embodiment of the present invention.

In step S201, the user inputs an English notation to which pronunciation symbols are to be given. In step S202, the notation processing unit 101 displays the English notation input in step S201. A of Fig. 9 shows a display example (note that Fig. 9 shows display examples in the direct input mode). In this example, assume that pronunciation symbols corresponding to an English notation "that" are input.

In step S203, the user presses a given key, and the key input processing unit 110 detects the key

pressed by the user.

The key input processing unit 110 checks in step S204 whether or not the key pressed by the user in step S203 is an "end key." If the pressed key is an "end key," the flow advances to step S223; otherwise, the flow advances to step S205.

The key input processing unit 110 checks in step S205 whether or not the key pressed by the user in step S203 is an "alphabet key." If the pressed key is an "alphabet key," the key input processing unit 110 stores that value in the input alphabet holding unit 111, and displays the input alphabet in an edit frame (A of Fig. 9). The flow then advances to step S206. If the pressed key is not an "alphabet key," the flow advances to step S212.

The pronunciation symbol candidate processing unit 102 checks in step S206 whether or not an alphabet is held in the input alphabet holding unit 111. If an alphabet is held, the flow advances to step S207; otherwise, the flow returns to step S203.

The pronunciation symbol candidate processing unit 102 determines with reference to the input mode holding unit 113 in step S207 whether or not the current input mode is the direct input mode. If the current input mode is the direct input mode, the flow advances to step S208; otherwise (i.e., the associative input mode), the flow advances to step S209.

If the current input mode is the direct input mode, the pronunciation symbol candidate processing unit 102 reads out, from the pronunciation symbol table 105, pronunciation symbol candidates corresponding to the alphabet held in the input alphabet holding unit 111 in step S208. For example, if the alphabet is "a," the corresponding pronunciation symbol candidates are "AA, AE, AH, AO, AW, AY." Note that pronunciation symbols of the English notation "that" in this example (Fig. 9) include a pronunciation symbol starting from alphabet "d," that starting from alphabet "a," and that starting from alphabet "t." Hence, the user inputs alphabet "d" initially, and "D, DH" are read out as candidates of pronunciation symbols that start from "d."

On the other hand, if the current input mode is the associative input mode, the pronunciation symbol candidate processing unit 102 reads out, from the associative pronunciation symbol table 105, pronunciation symbol candidates corresponding to the alphabet held in the input alphabet holding unit 111, and holds them in the pronunciation symbol candidate holding unit 103 in step S209. For example, when the alphabet is "a," corresponding pronunciation symbol candidates are "AA, AE, AH, AO, AW, AY, EH, ER, EY, IH, IY, OW." In case of the English notation "that" in this example (Fig. 9), the user inputs alphabet "t,"

and "CH, DH, SH, T, TH" are read out as pronunciation symbol candidates.

In step S210, the pronunciation symbol candidate processing unit 102 gives statistical values to the pronunciation symbol candidates held in the pronunciation symbol candidate holding unit 103 with reference to the pronunciation symbol statistical information 107. Furthermore, the unit 102 sorts the pronunciation symbol candidates in ascending order of statistical value.

In step S211, the pronunciation symbol candidate presentation unit 104 assigns image data to the pronunciation symbol candidates held in the pronunciation symbol candidate holding unit 103 with reference to the pronunciation symbol image data 108. Furthermore, the unit 104 presents the pronunciation symbol candidates assigned with the image data to the user. B of Fig. 9 shows a display example. In this case, pronunciation symbol candidates "D[d] DH[δ]" corresponding to user's input "d" are presented. Also, the first candidate "D[d]" is presented in an active state.

In this example, the unit 104 presents pronunciation symbol candidates assigned with the pronunciation symbol image data 108 to the user. Alternatively, the unit 104 may present pronunciation symbol candidates assigned with the pronunciation

symbol auxiliary data 109 to the user. In this case, "D[dee: D IY] DH[thee: DH IY]" are presented to the user.

The key input processing unit 110 checks in step S212 whether or not the key pressed by the user in step S203 is an "input mode change key." If the pressed key is an "input mode change key," the flow advances to step S213; otherwise, the flow advances to step S214.

In step S213, the input mode change unit 112 refers to the input mode held in the input mode holding unit 113. If the input mode is the "direct input mode" it is changed to the "associative input mode," or vice versa, and the flow advances to step S206.

The key input processing unit 110 checks in step S214 if the key pressed by the user in step S203 is a "select key." If the pressed key is a select key, the flow advances to step S215; otherwise, the flow advances to step S218.

The pronunciation symbol candidate presentation unit 104 checks in step S215 if pronunciation symbol candidates are presented to the user. If pronunciation symbol candidates are presented, the flow advances to step S216; otherwise, the flow returns to step S203.

In step S216, the pronunciation symbol presentation unit 104 changes an active one of the pronunciation symbol candidates presented to the user to the next candidate. The active candidate is, for

example, underlined. C of Fig. 9 shows an example.

In step S217, the pronunciation symbol speech generation unit 115 reads out speech data of the pronunciation symbol which is newly activated in step S216 from the phonemic symbol dictionary 116 and generates that speech data. The flow then returns to step S203.

The key input processing unit 110 checks in step S218 if the key pressed by the user in step S203 is an "enter key." If the pressed key is an "enter key," the flow advances to step S219; otherwise, the flow returns to step S203.

The pronunciation symbol candidate presentation unit 104 checks in step S219 if pronunciation symbol candidates are presented to the user. If pronunciation symbol candidates are presented, the flow advances to step S220; otherwise, the flow returns to step S203.

In step S220, the pronunciation symbol candidate presentation unit 104 presents the active pronunciation symbol in place of the alphabet in the edit frame. D of Fig. 9 shows an example.

In step S221, the pronunciation symbol candidate presentation unit 104 clears the presented candidates. E of Fig. 9 shows an example. The pronunciation symbol candidate processing unit 102 clears the pronunciation symbol candidates held in the pronunciation symbol candidate holding unit 103, and the flow advances to

step S222.

In step S222, the key input processing unit 110 clears the alphabet held in the input alphabet holding unit 111, and the flow returns to step S203. The aforementioned processes are repeated for the next pronunciation symbol (F of Fig. 9), thus finally inputting pronunciation symbols shown in G of Fig. 9.

In step S223, the edit result save unit 117 saves a pair of the input English notation and the edited pronunciation symbols in the edit result database 118.

As can be seen from the above description, according to this embodiment, in the direct input mode, the user need only input the first alphabet of a pronunciation symbol to display pronunciation symbols that start from the input alphabet and are sorted in descending order of predetermined probability of occurrence. Hence, compared to selection from an external character symbol table (about 40 symbols), the input efficiency can be greatly improved. In the associative input mode, pronunciation symbols when an alphabet forms a part of an arbitrary English notation are stored as associative pronunciation symbol information for respective alphabets. Every time the user inputs each alphabet that forms an English notation, pronunciation symbols corresponding to the input alphabet are displayed while being sorted in descending order of predetermined probability of

occurrence. Hence, compared to the conventional method (a method setting a pronunciation symbol in correspondence with one or two alphabets), the correspondence between alphabets and pronunciation symbols is clear, and an accurate input can be realized. As a result, pronunciation symbols can be efficiently and accurately input.

[Another Embodiment]

Note that the present invention may be applied to either a system constituted by a plurality of devices (e.g., a host computer, interface device, reader, printer, and the like), or an apparatus consisting of a single equipment (e.g., a copying machine, facsimile apparatus, or the like).

The objects of the present invention are also achieved by supplying a storage medium, which records a program code of a software program that can implement the functions of the above-mentioned embodiments to the system or apparatus, and reading out and executing the program code stored in the storage medium by a computer (or a CPU or MPU) of the system or apparatus.

In this case, the program code itself read out from the storage medium implements the functions of the above-mentioned embodiments, and the storage medium which stores the program code constitutes the present invention.

As the storage medium for supplying the program

code, for example, a floppy® disk, hard disk, optical disk, magneto-optical disk, CD-ROM, CD-R, magnetic tape, nonvolatile memory card, ROM, and the like may be used.

The functions of the above-mentioned embodiments may be implemented not only by executing the readout program code by the computer but also by some or all of actual processing operations executed by an OS (operating system) running on the computer on the basis of an instruction of the program code.

Furthermore, the functions of the above-mentioned embodiments may be implemented by some or all of actual processing operations executed by a CPU or the like arranged in a function extension board or a function extension unit, which is inserted in or connected to the computer, after the program code read out from the storage medium is written in a memory of the extension board or unit.

The present invention is not limited to the above embodiments and various changes and modifications can be made within the spirit and scope of the present invention. Therefore to apprise the public of the scope of the present invention, the following claims are made.